Helsingin yliopisto - Helsingfors universitet - University of Helsinki ID 2003-1997

Tiedekunta-Fakultet-Faculty	Laitos-Institution-Department
Valtiotieteellinen tiedekunta	Department of Economics

Tekijä-Författare-Author
Bäcklund, Galit

Työn nimi-Arbetets titel-Title
Is the Hypothesis of Homogeneity, in the Specification of Panel Data Models, a Dangerous or Harmless Simplification?

Oppiaine-Läroämne-Subject
Economics

Työn laji-Arbetets art-Level	Aika-Datum-Month and year	Sivumäärä-Sidantal-Number of pages
Master's thesis	2003-01-30	69

Tiivistelmä-Referat-Abstract
The hypothesis of homogeneity is a regression model simplification, which assumes that the slope coefficients of the regression equations are homogeneous hence, equal for all the agents (across sections). Although the hypothesis of homogeneity underlies the most widely applied panel data regression models, econometric literature is inconclusive about whether it is a dangerous or a harmless model simplification.

At first this thesis introduces the panel data concept and a profound panel data notation based on the forthcoming discussion paper by Professor Yrjö Vartia, where panel data set is presented in a geometric cubic form, the Panel Data Cube (PDC). Different types of data sets are also considered with respect to the PDC producing additional notations, Collapsed Panel Data Cubes (CPDC) and Conditional Panel Data Sections (CPDS). These pioneer notations are essential for understanding different panel data models.

Secondly, a survey of selected research papers that examine regression models that are based on the hypothesis of homogeneity ("pooled models") versus their counterparts is given. These research papers are classified according to the approach and the solutions they provide to the dilemma, leading me to detect a possible reason for the contradicting solutions in the literature.

Finally, an attempt is made to approach and solve the dilemma on a methodological level rather then the technical one literature provides us with. The developed theory classifies panel data regression models into three types with respect to the hypothesis, the Homogeneous Model (HOM), the Parallel Model (PM), and the Heterogeneous Model (HEM). The essence and the necessity of the hypothesis are examined through the three types of model and two modeling strategies. The thesis closes with a simplified proof, which illustrates that in general the hypothesis does causes regression distortion with group-regressions, but with panel data, when it is considered as a whole, these distortions are annihilated and no bias arise on average.

Avainsanat-Nyckelord-Keywords
hypothesis of homogeneity
panel data
regression model
pooled model
heterogeneous panels
heterogeneity

Säilytyspaikka-Förvaringsställe-Where deposited

Muita tietoja-Övriga uppgifter-Additional information